Thresholds in the resilience of modular social networks to invasion by defectors

Wechsler, Daniel ; Bascompte, Jordi

Abstract: Cooperative interactions constitute the backbone of many biological and social systems. Since cooperation is prone to exploitation, these systems must incorporate mechanisms that prevent the spreading of defective behaviors. One such mechanism is modularity, i.e., the tendency of a social network to be organized in modules, where individuals within a module tend to interact strongly among themselves while avoiding interacting with individuals from other modules. This structure allows cooperation to prevail by having modules of cooperative individuals with a limited exposure to defectors. To address the rate and shape of the effect of modularity on the resilience of cooperation, here we study a variant of the Prisoner's Dilemma on modular networks. Our simulations reveal a sharp transition between a resilient and a vulnerable regime as modularity exceeds a critical threshold. By using a simplified mathematical model, we show that the observed threshold is equivalent to the epidemic threshold found in a certain class of SIR models. This allows us to derive an explicit condition under which a cooperative society is expected to be resilient to invasive defectors.

DOI: https://doi.org/10.1016/j.jtbi.2018.10.018

Journal of Theoretical Biology 460 (2019) 56–63

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtb

Thresholds in the resilience of modular social networks to invasion by defectors

Daniel Wechsler*, Jordi Bascompte

Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, Zurich CH-8057, Switzerland

ARTICLE INFO

Article history:
Received 18 May 2018
Revised 4 October 2018
Accepted 5 October 2018
Available online 6 October 2018

Keywords:
Evolution of cooperation
Game theory
Non-linear transition
Complex networks

ABSTRACT

Cooperative interactions constitute the backbone of many biological and social systems. Since cooperation is prone to exploitation, these systems must incorporate mechanisms that prevent the spreading of defective behaviors. One such mechanism is modularity, i.e., the tendency of a social network to be organized in modules, where individuals within a module tend to interact strongly among themselves while avoiding interacting with individuals from other modules. This structure allows cooperation to prevail by having modules of cooperative individuals with a limited exposure to defectors. To address the rate and shape of the effect of modularity on the resilience of cooperation, here we study a variant of the Prisoner's Dilemma on modular networks. Our simulations reveal a sharp transition between a resilient and a vulnerable regime as modularity exceeds a critical threshold. By using a simplified mathematical model, we show that the observed threshold is equivalent to the epidemic threshold found in a certain class of SIR models. This allows us to derive an explicit condition under which a cooperative society is expected to be resilient to invasive defectors.

1. Introduction

The functioning of many biological and social systems relies on cooperative interactions between individuals. Striking examples of such systems include multicellular organisms, eusocial insects, or human societies. Since cooperation is prone to exploitation by defectors, it is a major scientific challenge to explain how such systems can evolve and prevail.

Among other mechanisms, a non-random social structure has been found to foster the evolutionary success of cooperative strategies (Nowak, 2006; Nowak and May, 1992). This finding arose largely from the study of models that extend classical game theory by explicitly incorporating a network that determines the contacts individuals can make (Lieberman et al., 2005; Nowak, 2006). In particular, the individuals occupy the nodes of the network and act either as cooperators or defectors. They then engage with their neighbors in games such as the Prisoners Dilemma, that capture the vulnerability of cooperators to exploitation by defectors. Evolution is incorporated by making the probability of reproduction dependent on the payoff individuals acquire from repeated rounds of the game. Studies based on such models further indicate that the precise structure of the contact network also matters. In particular, it was found that heterogeneous networks, which resemble those observed in real systems, are more favorable for cooperation than, for instance, regular networks (Santos et al., 2006).

Many real networks are highly modular. Modularity is the tendency of nodes to form distinct groups (modules) with connections running more likely between nodes within the same group than between nodes in different groups. Examples of real networks with a modular structure include metabolic networks (Ravasz, 2002), ecological communities (Grilli et al., 2016) or social networks (Girvan and Newman, 2002). It is widely believed, that a modular organization can generally increase the global resilience of a system to local perturbations (May, 1972). The intuitive argument is that a sparse inter module connectivity would buffer the spread of a perturbation across module boundaries (May, 1972; Stouffer and Bascompte, 2011). There are indeed theoretical predictions as well as a recent experimental result (Gilarranz et al., 2017) that suggest that a modular structure can indeed buffer the spreading of local perturbations.

Here, we investigate the potential of such a buffering effect in the context of the evolution of cooperation. In particular, we are interested in how modularity affects the resilience of a cooperative society against a collapse induced by a single invasive defector. In order to shed light on this question we combine simulations and analytical work using the model proposed in Santos et al. (2006).

* Corresponding author.
E-mail address: daniel.wechsler@ieu.uzh.ch (D. Wechsler).

https://doi.org/10.1016/j.jtbi.2018.10.018
0022-5193/

2. Methods

2.1. Game dynamics

We use the model proposed in Santos et al. (2006) to investigate how the resilience of a cooperative society against the invasion of a defective individual depends on the modularity of the underlying contact network. This contact network is assumed to be static and represents the social structure of the society. In particular, each individual occupies a node on the network and interacts only with its direct neighbors. The evolutionary process advances in discrete time steps (i.e., iterations), with a single time step involving two phases. In the first phase, the fitness of each individual is evaluated. An individual's fitness is taken to be the sum of the payoff it acquires from a single round of the Prisoner's Dilemma (PD) game against each of its direct neighbors. In the PD, each individual acts according to its innate strategy as either a cooperator (C) or defector (D). The game parametrization we use depends on a single parameter $b \in [1, 2]$, that determines the payoff advantage of a defector when playing against a cooperator (cheating advantage). The payoff matrix is shown in the table below with the values indicating the payoff of the row player given its own strategy and that of its opponent (columns):

	C	D
C	1	0
D	b	0

In the second phase, each individual competes with a randomly chosen neighbor for the node it is currently occupying. If the chosen neighbor accumulated a higher payoff, an offspring of it replaces the focal individual with a probability proportional to the payoff difference between the competing individuals. This probability is given by the following equation:

$$q = max\left[\frac{z_j - z_i}{b\, k_{max}},\ 0\right], \tag{1}$$

where z_i indicates the payoff sum of the focal individual and z_j represents the payoff sum of the randomly chosen neighbor. The factor k_{max} is the largest degree of the two nodes and assures that q is in the unit interval. All network nodes are updated synchronously (i.e., in a single time step all individuals are updated) and a simulation ends after a predefined number of time steps T (see supplementary material for a pseudo code of the game dynamics).

2.2. Modular network generation

We consider networks with m statistically identical modules. That is, we assume that each module has the same number of nodes n, the same number of internal connections L_{in} (i.e. connections within modules), and the same number of outgoing connections L_{out} (i.e. connections between modules). Further, we require that the number of connections between each pair of modules is the same and therefore given by $L_{out}/(m-1)$. To quantify the modularity of such a network we use the measure proposed by Newman and Girvan (2004) which, for the subset of network structures considered here, can be written as Trajanovski et al. (2013):

$$Q = 1 - \frac{1}{m} - \frac{L_{out}}{2L_{in} + L_{out}} \tag{2}$$

This measure captures the tendency of nodes to be preferably connected to nodes within the same module. It can take values in the interval $[-\frac{1}{2}, 1 - \frac{1}{m}]$, with the maximal value taken for a network with m isolated modules.

Given a number of modules m, nodes per module n, and total number of links L, a network of a desired modularity can be constructed by calculating the right proportion of within module L_{in} and between module L_{out} links using equation (2) and the fact that $L = mL_{in} + \frac{1}{2}mL_{out}$. We followed the following two step procedure in order to assemble the network (Decelle et al., 2011; Karrer and Newman, 2011) (see supplementary material for a pseudo-code):

(i) For each of the m modules, we created the internal connections by placing a link between L_{in} randomly chosen node pairs. We required that each of the n nodes has at least one internal connection and that there can not be more than one connection between any node pair. Hence, each module by itself is an Erdös-Rényi random graph with L_{in} edges (Erdös and Rényi, 1960).

(ii) Between each pair of modules we establish $L_{out}/(m-1)$ links with the end points chosen at random. Also, in this step we do not allow for multiple edges between the same two nodes.

By keeping L constant and varying only the ratio of L_{in} to L_{out}, this procedure allows us to generate networks with different values of modularity but constant average node degree and statistically invariant degree distributions. There are however certain constraints on the parameter combinations for which a network can be realized. Most important, for a network with modules of size n, the maximum number of internal links L_{in} is $n(n-1)/2$, which is the case when each node pair in a module is connected, i.e., modules are cliques. Although the minimum number of internal links required for n nodes to form a connected component is $n-1$, we considered only networks with $L_{in} > \frac{1}{2}\log(n-1)$, which designates the boundary above which random networks form almost surely a single connected component (Erdös and Rényi, 1960). Further, we only considered networks for which the partitioning of nodes into modules, that we impose, is the one that maximizes Q. We controlled for that by comparing the modularity value of our partitioning with the modularity value of the best partitioning found by commonly used modularity detection algorithms (Newman, 2006; Reichardt and Bornholdt, 2006).

2.3. Simulations

We started each simulation from a single defector in a network of otherwise only cooperators. The model was then iterated for a predefined number of time steps T. The duration T was chosen such that it was long enough for the simulations to reach an equilibrium. Since this value depends on system size we indicate it for each computational experiment separately in the main text. As will be explained in the results section, defectors have a competitive advantage within modules and therefore, upon invasion of a module, they replace the majority of the resident cooperators. Thus, at equilibrium we find modules to be either occupied by only cooperators (those that resisted invasion) or by a majority of defectors. As our resilience measure we took the fraction of modules that resisted defector invasion. For a single simulation we computed this quantity from the state at time step T. To account for the stochasticity of the model, we calculated the final data for a particular parameter combination as an average from a large number of simulations using a newly generated (i.e., randomized) network each time. The actual number of simulations used per parameter combination depends on the computational experiment and is indicated at the respective position in the main text.

3. Results

Fig. 1 shows the fraction of modules able to resist defector invasion as a function of modularity for different values of the cheating advantage b.

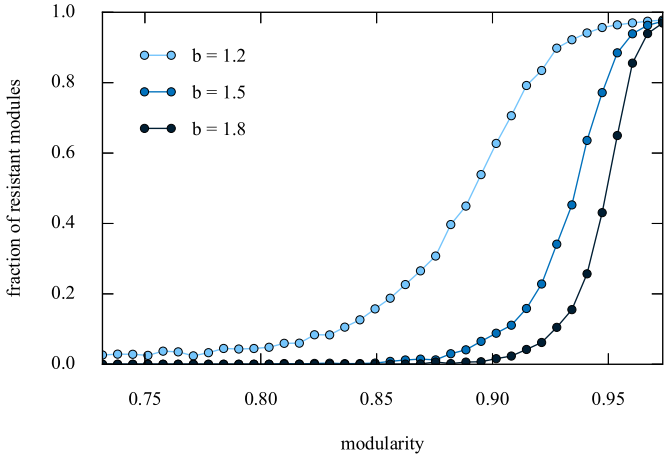

Fig. 1. Fraction of modules able to resist defector invasion, as a function of modularity Q, for different values of the cheating advantage b, and average degree $\bar{k} = 100$ (number of modules $m = 50$, and nodes per module $n = 150$). The value of each data point is the average from 600 simulations ($T = 1200$).

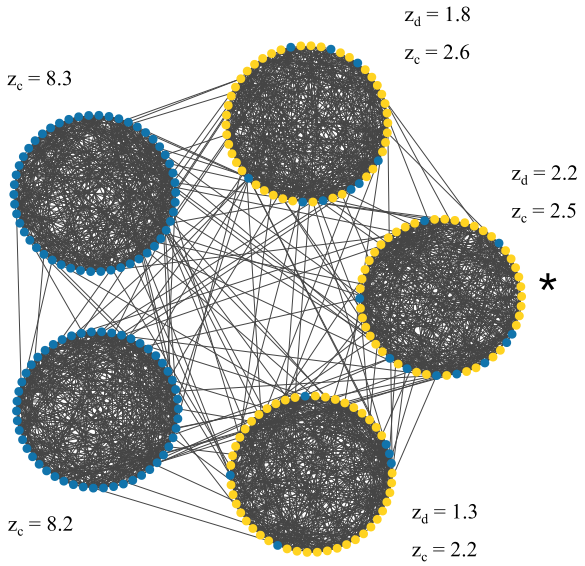

Fig. 2. Typical equilibrium? configuration of cooperators (blue) and defectors (yellow) on the $N = 250$ nodes of a highly modular network with $m = 5$ modules. The numbers indicate the average payoff of defectors ($\bar{z}_d$) and cooperators ($\bar{z}_c$) within each module and the asterisk (*) indicates the module where the first defector originated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The effect of modularity on the fraction of resistant modules is highly non-linear. In particular, we observe a sharp transition from a vulnerable to a resilient regime as modularity is increased. In the vulnerable regime, all modules fall victim to defectors, while in the resilient regime, defectors are constrained to the module where they initially appear.

Fig. 2 shows a typical equilibrium state along the transition, when neither cooperators nor defectors dominate. For illustration purposes, we used a network with only five modules and grouped the 250 nodes accordingly. In this case, defectors (yellow nodes) originate from the right-most module (indicated by the asterisk) and managed to expand into two additional modules. We will refer to modules that fall victim to defectors as *defector modules* and those that are only occupied by cooperators (blue nodes) as *cooperator modules*.

The numbers next to each module indicate the average payoff of defectors ($\bar{z}_d$) and cooperators ($\bar{z}_c$) within each module. Since

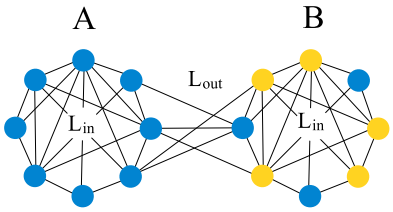

Fig. 3. Illustration of a two module network with L_{in} the number of edges within each module and L_{out} the number of edges between modules. Both modules are internally fully connected and hence the number of nodes per module n constraints the number of internal edges to $L_{in} = (n^2 - n)/2$. Modularity Q_2 of this network depends only on L_{out} and has its maximum at $\frac{1}{2}$.

defectors profit only from a few interactions with cooperators, their average payoff is much lower than the one of cooperators on cooperator modules. This payoff difference prevents defectors from expanding further into the two remaining cooperator modules. Likewise, cooperators can not take over defector modules. The insular cooperators, originating from adjacent cooperator modules, can not expand further because their offspring is exploited by the defector dominated neighborhood.

Within a module, defectors have a competitive advantage over cooperators. Thus, if defectors can take over a node in a cooperator module, the majority of the resident cooperators will be replaced. This local advantage of defectors, however, counteracts their prospect of expanding further into adjacent modules, because the co-resident cooperators are their primary payoff source.

As it will be shown, this density dependent effect determines the probability with which defectors can overcome module boundaries and consequently the expected fraction of modules they can invade. In what follows, we derive a mathematical expression for this *invasion probability* for the simplified two module system depicted in Fig. 3. We then show that knowing the invasion probability for a two module system is enough to predict the fraction of modules invaded in a multi-module network. In a last step, we show that even though our expression for the invasion probability is derived for the special case in which modules are cliques, predictions based on it also hold if this constraint is relaxed.

3.1. Invasion probability for a two module system

For the two module system in Fig. 3, we define the invasion probability p_2 as the expected probability with which defectors (yellow), when expanding from a single randomly chosen node in module B, manage to invade the cooperator module A (take over a node in A). We will derive an expression for p_2 as a function of modularity Q, the cheating advantage b and the number of nodes per module n under the simplifying assumption that both modules are cliques, i.e., within a module each node is connected to each other node. With this constraint the modularity of the network is only a function of L_{out} (since $m = 2$ and $L_{in} = (n^2 - n)/2$). We will use the symbol Q_2 to refer to the modularity of this type of two module network.

Consider the strategy update of a single node in module A. For the present cooperator to be replaced by a defector from module B three conditions need to be met:

First, the cooperator must compete against an individual on a node in module B. On average this happens with probability $k_{out}/(k_{in} + k_{out})$, where k_{in} is the expected number of connections the cooperator node has to other nodes in module A and k_{out} is the expected number of connections to nodes from module B. Since $k_{out} = L_{out}/n$ and $k_{in} = 2L_{in}/n$, the probability for a cooperator to compete against a node from module B is given by $\alpha = L_{out}/(2L_{in} + L_{out})$, which corresponds to the last term in equation (2). Independent of m, this term is the probability of an individual

to engage in an inter module interaction. We can also write this probability in terms of the modularity of the two module network as $\frac{1}{2} - Q_2$.

Second, the selected node from module B must be occupied by a defector. If we assume that nodes in module B that have a connection to module A are equally likely to be occupied by defectors and cooperators, then this probability is equal to the fraction of defectors d in module B.

Third, if indeed a defector occupies the chosen node, that defector must "beat" the resident cooperator. The expected probability of this to happen can be computed by applying the expected payoff of a cooperator in module A (z_c) and the expected defector payoff (z_d) to formula (1). The expected payoff of a cooperator in module A is $z_c = k_{in} + (1-d)k_{out}$ and that of a defector is $z_d = bk_{in}(1-d) + bk_{out}$. Now we make the simplifying assumption that individuals receive only a payoff from interactions with nodes in the same module. Hence, the last terms in the sums in the formulas for z_c and z_d are not taken into account. Given these three considerations, the probability r_2 for a cooperator in module A to be replaced by a defector during a single update, given the fraction of defectors d in module A, the degree of modularity Q_2, and the cheating advantage b, can be written as:

$$r_2(d) = d\,(\frac{1}{2} - Q_2)\; max\Big[1 - \frac{1}{b} - d, 0\Big]. \tag{3}$$

We wrote r_2 as a function of d because in contrast to Q_2 and b, d is a dynamic quantity (see supplementary material for more details on the derivation of Eq. 3).

Given a time series $d_0, d_1, \ldots d_T$ of defector frequencies in module B, we can calculate the probability with which defectors manage to invade module A up to time step T as

$$P_n(T) = 1 - \prod_{t=0}^{T}\Big[1 - r_2(d_t)\Big]^n\,, \tag{4}$$

where n is the number of nodes per module. This accounts for the fact that in a single iteration each node in the cooperator module could potentially be taken over by a defector.

The change in the frequency of defectors in module B in a single iteration, given the current frequency d, can be approximated as:

$$\delta(d) = d\,(1-d)^2\,(1 - \frac{1}{b})\,, \tag{5}$$

This function is approximate in the sense that, like equation (3), it ignores the payoff from interactions with module A and assumes n>>1. It is derived in the supplementary material along the lines of equation (3) as the product of i) the probability of a cooperator in module B to interact with a defector (in module B), ii) the probability of the defector replacing the cooperator, and iii) the current cooperator frequency. For $d \in (0,\ 1)$ and $b \in (1,\ 2]$, $\delta(d)$ is always positive, reflecting the fact that in a homogeneous society (fully connected module), defectors always out compete cooperators. Thus, a time series obtained by iterating $d_t = d_t + \delta(d_t)$, from an initial frequency d_0, follows a logistic growth curve with defectors always reaching fixation (within module B).

The black line in Fig. 4 shows such a growth curve when starting the iteration from a single defector $d_0 = 1/n$. The blue curve is the normalized payoff difference between defectors in module B and cooperators in module A. It corresponds to the first argument of the max function in equation (3) and thus determines the probability with which a cooperator in module A is replaced by a defector upon an interaction (negative parts not shown). As the number of defectors increases, this quantity decreases until it reaches zero at $d^* = 1 - \frac{1}{b}$. Once this threshold density is crossed, defectors are not anymore able to invade module A. This is indicated by the red curve, the probability with which defectors can invade module A in a given time step. It corresponds essentially to $r_2(d)$ but

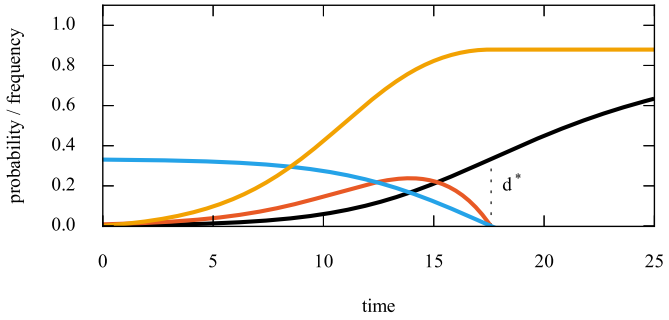

Fig. 4. Time evolution of the frequency of defectors within a module (black), probability of a defector in module B replacing a cooperator in module A upon an interaction (blue), per iteration invasion probability (red) and its temporal accumulation (yellow). The parameters used are $n = 400$, $b = 1.5$ and $Q = 0.476$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

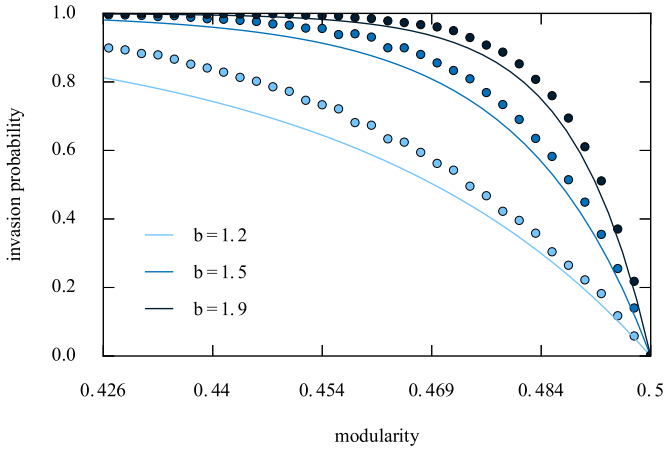

Fig. 5. Invasion probability p_2 as a function of the modularity Q_2 of a two module system ($m = 2$) with $n = 250$ nodes per module, computed using equation (6) for different cheating advantages b (lines). The dots represent results from simulations with the corresponding parameter values. For each Q_2, b combination we performed 5500 simulations for $T = 1400$ time steps and calculated the invasion probability as the fraction of simulations in which defectors could invade module A.

takes into account that there are n nodes that can potentially be taken over during a single iteration. The yellow curve corresponds to $P_n(t)$, the probability that defectors manage to invade module A up to time t. The asymptotic value of this curve, which is reached when $d(t) = d^*$, is the invasion probability p_2.

As explained in the supplementary material, by considering the continuous time limit case, p_2 can be expressed in terms of the following equation:

$$p_2 = 1 - e^{\int_{\frac{1}{n}}^{d^*} \frac{n}{\delta(d)} \log\left[1 - r_2(d)\right] dd} \tag{6}$$

The lines in Fig. 5 show p_2 as a function of Q_2 for different values of the cheating advantage (using numerical integration of equation (6)). The dots are simulation results computed as the fraction of 5500 simulations per data point in which defectors managed to invade module A when starting from a single defector in module B. Apparently, there is a good agreement between the predictions and the simulations. The tendency of a slight underestimation of the true invasion probability is most likely due to neglecting payoff from inter module interactions in equations (3) and (5).

3.2. Extension to multi-module system

After having derived an expression for the invasion probability for a two module system, we will now use it to predict the fraction

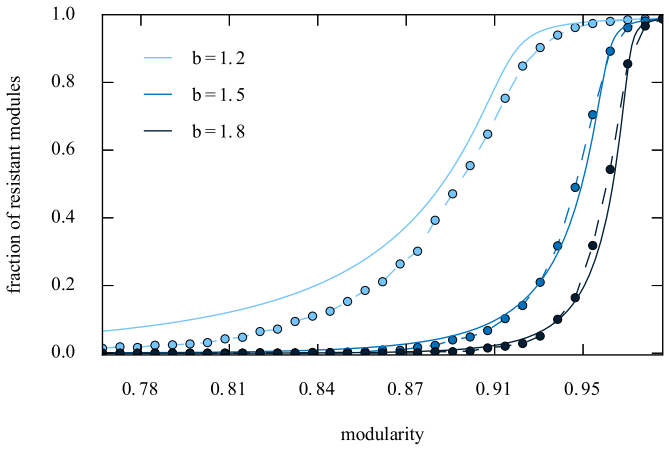

Fig. 6. Fraction of modules resisting defector invasion calculated as the asymptotic value of C_s divided by m (lines) and the same quantity computed as the average from 1000 simulations per data point and $T = 1800$.

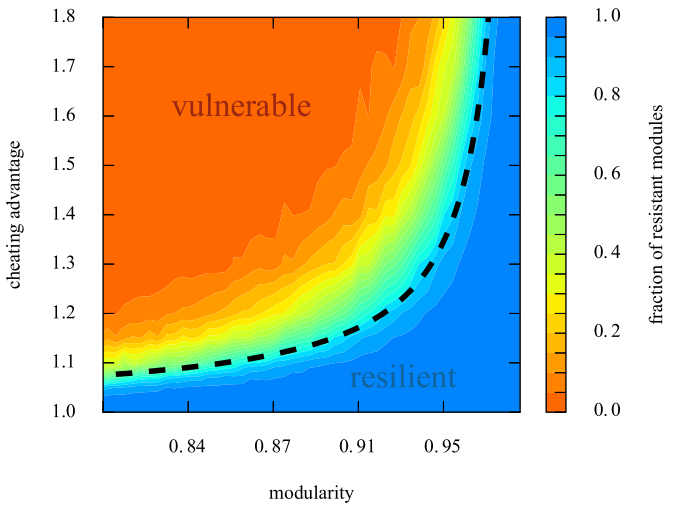

Fig. 7. Critical line (black) through the parameter space given by the network modularity Q and the cheating advantage b separates the vulnerable from the resilient regime (for fixed $n = 150$ and $m = 100$) as predicted by resilience condition (9) . The color indicates the fraction of modules resisting defector invasion obtained as an average from 200 to 1000 simulations per $Q - b$ combination ($T = 1800$).

of modules able to resist defector invasion in a multi-module network. For that, we define a compartment model that describes the system at the level of module states rather than node states. This model can be expressed in terms of the following three coupled difference equations:

$$\begin{aligned} U_{s+1} &= C_s(1-(1-p_m)^{U_s}) \\ C_{s+1} &= C_s - C_s(1-(1-p_m)^{U_s}) \\ D_{s+1} &= D_s + U_s \end{aligned} \tag{7}$$

The state variables C, D and U represent the number of cooperator, defector and transient modules in the network at time s. Transient modules are modules which are invaded by defectors, but with defector frequency not yet beyond the threshold value d^* (see Fig. 4). The model assumes that a unit of time s corresponds to the duration of the transient phase and hence, every transient module in time step s becomes a defector module at $s+1$ (third equation). Further, every transient module at time s can turn every cooperator module with probability p_m in a transient module at time $s+1$ (first and second equation). This reflects defectors invading cooperator modules from transient modules. The parameter p_m is the invasion probability specific for a network with m modules. For a pair of modules, p_m is the probability that defectors, when originating from a single individual in one of the two modules, can invade the other module via the direct links between the two modules. As shown in appendix B, an expression for p_m follows directly from equation (6) by taking into account that in an m module network there are only $L_{out}/(m-1)$ links between any module pair. The number of cooperator modules predicted to resist defector invasion can be obtained as the asymptotic value of C_s when iterating equations (7) from initial condition $C_0 = m-1$, $D_0 = 0$ and $U_0 = 1$.

The solid lines in Fig. 6 show this quantity as a function of modularity, for different values of b ($m = 100$, $n = 150$). The compartment model captures the transition observed in the individual based model. For large b there is also a good quantitative agreement, as seen from the small deviations to simulation results (dots).

In appendix A we show that for large m (or equivalently small p_m), our compartment model is equivalent to the Kermack-McKendrick model (Kermack and McKendrick, 1927). This is a variant of an SIR model that describes the spreading of an infectious disease in a homogeneous population. This equivalence allows us to adopt results derived for the SIR model to our system. In particular, the epidemic threshold, which cleaves the parameter space in a region where a disease outbreak is limited to a fraction of people and a region where it is expected to pervade the entire population (epidemic). A disease is below the epidemic threshold if the number of secondary infections caused by a single primary infection is less than one. As shown also in appendix A, the analogous of this threshold in our system can be written in terms of the following inequality:

$$p_m\ (m-1) < 1 \tag{8}$$

If the inequality holds, the system is resilient to defector invasion. This is the case if the expected number of modules invaded directly from the module where the first defector occurs is less than one. As explained in appendix B, for networks with large modules ($n \gg 1$) and a large number of modules ($m \gg 1$), resilience condition (8) can be rewritten in terms of modularity, module size, number of modules and cheating advantage as:

$$Q > 1 - \frac{1}{m} - \frac{1}{\frac{n}{2}(b-1)}, \tag{9}$$

thus, providing a threshold for the spreading of a defector that is explicitly related to the modularity of the network. Whenever the level of modularity of the social network is higher than the threshold in Eq. (9), cooperation will be stable in the face of defectors. The dotted line in Fig. 7 indicates the combinations of $Q - b$ values for which the two sides of inequality (9) are equal. It designates the boundary between the resilient (lower left) and the vulnerable (upper right) regimes through the $Q - b$ parameter space. The colors represent the fraction of modules resisting defector invasion computed from simulations. The predicted critical line quite well aligns with the threshold observed in simulations, indicating that inequality 9 seems suitable as a resilience condition.

While it is convenient to express the resilience condition in terms of modularity, it blurs the fact that the number of modules does not affect resilience. In particular, both the constant one and the term $1/m$ appear on both sides of inequality (9) and it can thus be rewritten as:

$$\alpha\ n\ \frac{1}{2}(b-1) < 1 \tag{10}$$

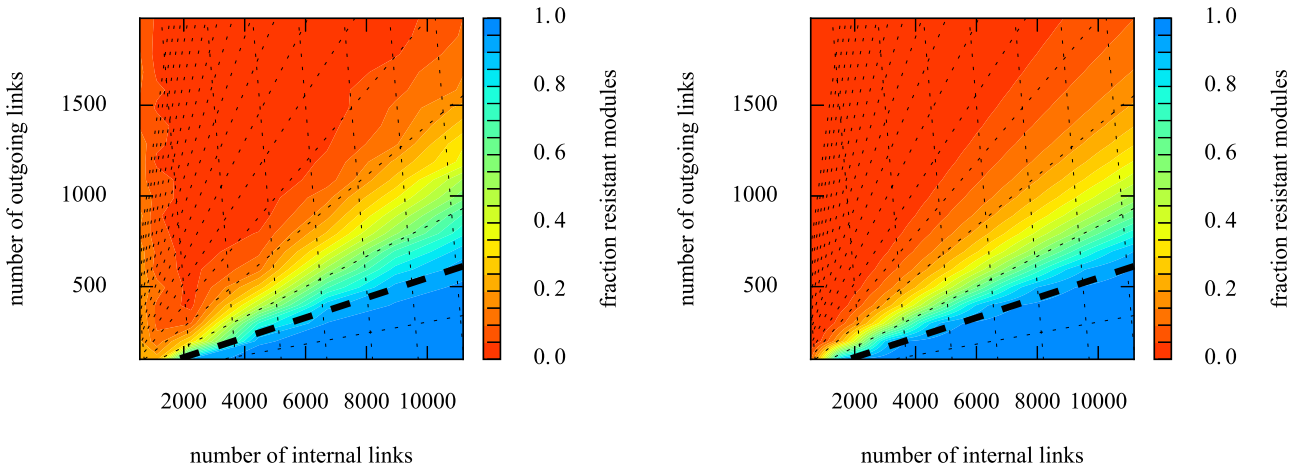

Fig. 8. Fraction of cooperator modules resisting defector invasion in a network with $m = 100$ modules and $n = 150$ nodes per module and $b = 1.5$. The critical line is indicated by the bold dashed line. The fine lines originating in the bottom left corner are isoclines of constant α. The lines running from top to bottom are isoclines of constant average degree z.

3.3. Non-homogeneous modules

Our formula for the invasion probability r_2 (and p_m) was derived based on the assumption of fully connected modules. Fig. 8 shows a comparison between simulations (left) and analytical predictions (right) if this assumption is relaxed.

In both plots the colors indicate the fraction of resistant modules as a function of the fraction of interactions within (L_{in}) and between (L_{out}) modules. For the simulations the values were computed as an average over 200 simulations per data point ($T = 1800$). For the analytical results we used, as for Fig. 6, the time asymptotic value of C_s (divided by m). The largest considered L_{in} value designates the fully connected limit (right borders) and the smallest value is where only around 5 percent of the possible internal links are realized.

As noted, both the course of the critical line (bold black) as well as the shape of the transition predicted by the compartment model is in good agreement with the simulation results. In the supplementary material we provide additional plots for other parameter values which convey a similar picture.

4. Discussion

We showed that a modular network structure can prevent the spreading of defectors in a cooperative society. This protective effect is, however, a rather all or nothing property that sets in only after modularity exceeds a critical threshold value. We found that the crucial quantity governing this abrupt transition between a vulnerable and a resilient regime is the probability with which defectors can invade adjacent cooperator modules before reaching a limiting local density. For a simplified two module system, we could derive a formula for this invasion probability. Moreover, it turned out that the two module case approximates well the invasion probability between pairs of modules in a multi-module system. This allowed us to use a compartment model, parameterized with the invasion probability, to predict the observed transition. Since the compartment model is equivalent to a well studied SIR model, we could adopt the condition derived for the epidemic threshold to our system. This led to an expression for the condition under which a cooperative society is resilient against defector invasion. This condition depends on modularity, module size, and the cheating advantage defectors have over cooperators.

Beyond supporting previous findings that a modular network structure protects cooperation (Fotouhi et al., 2018; Gianetto and Heydari, 2015; Lozano et al., 2008; Marcoux and Lusseau, 2013), our results also shed light on the type of relationship between modularity and resilience, as well as on the mechanisms that give rise to it. This understanding might help in choosing the appropriate action when trying to increase the resilience of cooperative societies (e.g. reducing inter module connectivity). Of value is for example the insight that the effectiveness with which a certain system modification increases resilience crucially depends on the current configuration of the system. This is best illustrated in Fig. 7, where it is apparent that depending on where in the $Q - b$ parameter space a system resides, it might be more effective to either lower the cheating advantage or increase modularity to avoid the vulnerable regime. Hence, based on our resilience condition one could, in principle, calculate the most effective measures to increase the resilience of a particular cooperative society.

The fact that our results were obtained for only a subset of conceivable structural and dynamical setups, puts however limits on the generality of our conclusions. Concerning structure, we only considered networks where all modules have the same size and internal connectivity and where the same number of edges connect each pair of modules. A convenience of this is that there is a unique invasion probability suitable for all pairs of modules. However, once heterogeneity is introduced among modules, by for instance varying the number of outgoing connections, we would have to deal with multiple invasion probabilities potentially two for each pair of modules. How the accuracy of predictions, based on an average invasion probability, would decline with the increase in heterogeneity is not clear. From studies on spreading processes it is however known that deviations from the homogeneous limit can result in substantial differences in, for example, the actual values of the critical threshold (Newman, 2002). Concerning the dynamical aspects, the generality of our results is limited in that we considered only a particular type of strategy update scheme. There are however multiple alternative schemes commonly used in studies on cooperation on networks (e.g. birth-death and death-birth update). Comparisons between different schemes revealed that the choice of the update scheme can affect whether cooperation is favored or not, on a particular network structure (Ohtsuki et al., 2006). Hence, while alternative update schemes would certainly change the results quantitatively, it might also be that certain schemes would produce qualitatively different results.

However, the convergence of results between our study and studies that used different setups (Fotouhi et al., 2018; Jiang and Perc, 2013) makes us confident that our results are not just a peculiarity of the model chosen here.

Another insight worth mentioning is concerned with the suitability of the modularity measure Q as a predictor for the response of a modular network to defector invasion. Inequality (10) states that there are two basic topological features determining resilience. First, the proportion of internal to external connections a module has (L_{in}/L_{out}), which determines the probability α with which an individual engages in an interaction with a node from another module. Second, the number of nodes per module n. From these two features, the measure Q only accounts for the interaction probability, but not for module size. Consequently, adding nodes to modules, while maintaining the proportion of internal to external connections, lowers resilience, but does not affect modularity. In addition, Q includes the number of modules m in its definition, which does not affect resilience (for $m \gg 1$). As a consequence, one can not increase resilience by increasing modularity through the inclusion of additional modules. Based on these observations, we conclude that the modularity measure Q, as defined in Newman and Girvan (2004), is not an ideal predictor for the response of a modular network to defector invasion. The more informative quantity turns out to be the product of interaction probability and number of nodes per module. Interpreted from a dynamical point of view, this quantity is the rate at which external interactions occur, i.e., the expected number of interactions all individuals of a module have with individuals from other modules per unit time.

In summary, our study is in line with the hypothesis that a modular organization increases the resilience of networked systems (see however Sah et al. (2017)). It however exemplifies that the relationships between modularity and resilience can be nonlinear. A pattern also revealed by recent theoretical studies investigating the effect of network modularity on the fixation probability of cooperators within a defective society (Fotouhi et al., 2018; Jiang and Perc, 2013). Further work needs to clarify the implications of this non-linearity in for example the ability to anticipate shifts in the resilience of evolving systems.

Acknowledgments

We thank A. Ferrera for useful suggestions at different stages of the study, Raul Ortega for technical support, and M. A. Fortuna and M. Barbour for useful comments on the final manuscript. This work was supported by the European Research Council through an Advanced Grant and by the Swiss National Science Foundation (grant 31003A_160671 to J.B.).

Appendix A

A1. A: Equivalence of the compartment model and Kermack-McKendrick model

Here we show that the compartment model defined by the difference equation system (7) is, for small p_m, equivalent to the Kermack-McKendrick model (Kermack and McKendrick, 1927). This basic variant of an SIR model is defined by the following system of coupled differential equations:

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI\\ \frac{dI}{dt} &= \beta SI - \gamma I\\ \frac{dR}{dt} &= \gamma I.\end{aligned} \tag{A.1}$$

The previous system models the spreading of an infectious disease in a homogeneous population, with S being the number of susceptible people, I being the number of infected people and R being the number of people who recovered from the disease and developed immunity. The two parameters are the infection rate β and the recovery rate γ.

The equivalence to our compartment model is best seen by considering the continuous limit of the difference equation system (7). Note, in contrast to the main text, here we use t instead of s to refer to time in the compartment model. In the continuous limit, the temporal change in the number of cooperator modules C is given by (in the last step we applied L'Hospital's rule):

$$\begin{aligned}\frac{dC}{dt} &= \lim_{\Delta t\to 0}\frac{C_{t+\Delta t}-C_t}{\Delta t}\\ &= \lim_{\Delta t\to 0}\frac{-C_t(1-(1-p_m)^{U_t\Delta t})}{\Delta t}\\ &= C_tU_t\log(1-p_m)\end{aligned}$$

The continuous time equation for the change in the number of transient modules U follows directly from the previous equation as:

$$\frac{dU}{dt} = -C_tU_t\log(1-p_m) - U_t$$

Further, the continuous time equation for the frequency of defector modules (D) is:

$$\frac{dD}{dt} = U_t$$

Since for $p_m \ll 1$ $\log(1-p_m) \approx -p_m$, we can write the change of the number of cooperator modules as:

$$\begin{aligned}\frac{dC}{dt} &\approx -p_mCU\\ \frac{dU}{dt} &\approx p_mCU - U\\ \frac{dD}{dt} &= U\end{aligned} \tag{A.2}$$

From Eq. (A.2), the equivalence to Eq. (A.1) is apparent by the correspondence $S \to C$, $I \to U$, $R \to D$ between state variables and the correspondence $\beta \to p_m$, $\gamma \to 1$ between parameters.

The crucial quantity for the epidemic threshold in the Kermack-McKendirck model is the basic reproductive number R_0 Weisstein. It is defined as the number of secondary infections caused by a single infected individual introduced into a population of otherwise susceptible individuals (note that the original naming is a bit misleading since R_0 has nothing to do with R the number of recovered individuals). Formally, the reproductive number can be written as:

$$R_0 = \frac{\beta S}{\gamma}$$

If $R_0 > 1$ a disease is expected to spread because an individual who carries it infects more than one susceptible. In contrast, if $R_0 < 1$ a disease outbreak is expected to wane. Given the equivalence of this SIR model to our compartment model, the basic reproduction number translates to the expected number of cooperator modules invaded (directly) from a single transient module. This quantity can be written as $p_m\,C$, with p_m being the invasion probability in an m module network. Stated as a condition that, if fulfilled, implies resilience to defector invasion, and also assuming that initially $C_0 = m-1$ we have:

$$p_m\,(m-1) < 1$$

Apparently, at the critical threshold, p_m is small if the number of modules m is large. Hence, we expect results derived for the Kermack-McKendrick model to also hold for our system if the number of modules is sufficiently large.

A2. B: Derivation of the resilience condition

Based on the equivalence of our compartment model (equations (7)) to the Kermack-McKendirck SIR model, we could derive inequality (8) as a condition that, if satisfied, implies the resilience of a cooperative society to defector invasion (see appendix A). In what follows, we restate this condition in terms of the basic model parameters Q, b, n, and m. We do so by using an approximation of p_m that holds if both m and n are large.

A formula for p_m follows directly from that for r_2, by incorporating the fact that in an m module network the probability of an individual to interact with an individual from another module is given by $\alpha/(m-1)$ (with $\alpha = L_{out}/(2L_{in}+L_{out})$ as defined in Section 3.1). Hence, considering a cooperator module and a module occupied by a fraction d of defectors, the probability of a cooperator on the cooperator module to be replaced by a defector from the other module during a single strategy update is given by:

$$r_m(d) = d\,\frac{\alpha}{(m-1)}\left[1-d-\frac{1}{b}\right] \tag{B.1}$$

This formula corresponds to $r_2(d)$ scaled by $1/(m-1)$ and written in terms of α instead of modularity, which is more convenient for the following computations. By replacing $r_2(d)$ by $r_m(d)$ in equation (6) we can write p_m as:

$$p_m = 1 - e^{\int_{\frac{1}{n}}^{1-\frac{1}{b}} \frac{n}{\delta(d)}\log[1-r_m(d)]dd} \tag{B.2}$$

Instead of solving Eq. (B.2) directly, we approximate p_m through its linearization at $\alpha = 0$:

$$p_m \approx \alpha\,\frac{dp_m}{d\alpha}\bigg|_{\alpha=0} = \alpha\,\frac{n}{(m-1)\left(1-\frac{1}{b}\right)}\left[\log\left(b-\frac{b}{n}\right)+\frac{1}{b-\frac{b}{n}}-1\right]$$

We show the computation of the derivative as well as the validity of the approximation for systems with $m \gg 1$ in the supplementary material. If we further assume that the number of nodes per module n is also large, then the above expression simplifies to:

$$p_m \approx \frac{\alpha}{(m-1)}\,n\left[\frac{\log(b)}{(1-\frac{1}{b})}-1\right]$$

Plugging this approximation of p_m into the resilience condition (8) gives:

$$\alpha\,n\left[\frac{\log(b)}{(1-\frac{1}{b})}-1\right] < 1 \tag{B.3}$$

The term $(m-1)$ canceled out, indicating that the number of modules does not affect the critical threshold. This is because near the threshold the expected number of modules invaded from a transient module is approximately equal to the expected number of nodes invaded from that module (which is what the left side of the inequality reflects). This holds because we assume m >> 1, which implies that around the threshold the expected number of modules being invaded is small compared to the actual number of modules.

We can rewrite inequality (B.3) in terms of modularity using equation (2), which reintroduces m into the condition (but with no effect since it appears on both sides of the inequality).

$$Q > 1 - \frac{1}{m} - \frac{1}{n\left[\frac{\log(b)}{(1-\frac{1}{b})}-1\right]}. \tag{B.4}$$

Since for $b \in [1, 2]$ the term $\log(b)/(1-\frac{1}{b})$ is well approximated by $\frac{1}{2}(b+1)$, the two versions of the resilience condition can be expressed in an even simpler form as:

$$\alpha\,n\,\frac{1}{2}(b-1) < 1 \tag{B.5}$$

$$Q > 1 - \frac{1}{m} - \frac{1}{\frac{n}{2}(b-1)} \tag{B.6}$$

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jtbi.2018.10.018

References

Decelle, A., Krzakala, F., Moore, C., Zdeborová, L., 2011. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. Phys. Rev. E. 84 (6), 066106. doi:10.1103/physreve.84.066106.

Erdös, P., Rényi, A., 1960. On the evolution of random graphs. Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei 5, 17–61.

Fotouhi, B., Momeni, N., Allen, B., Nowak, M.A., 2018. Conjoining uncooperative societies facilitates evolution of cooperation. Nature Human Behav. 2 (7), 492–499. doi:10.1038/s41562-018-0368-6.

Fotouhi, B., Momeni, N., Allen, B., Nowak, M. A., 2018. Evolution of Cooperation on Stochastic Block Models1807.03093.

Gianetto, D.A., Heydari, B., 2015. Network modularity is essential for evolution of cooperation under uncertainty. Sci. Rep. 5 (9340). doi:10.1038/srep09340.

Gilarranz, L.J., Rayfield, B., Liñán-Cembrano, G., Bascompte, J., Gonzalez, A., 2017. Effects of network modularity on the spread of perturbation impact in experimental metapopulations. Science 357 (6347), 199–201. doi:10.1126/science.aal4122.

Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. Proceed. Nation. Acad. Sci. 99 (12), 7821–7826. doi:10.1073/pnas.122653799.

Grilli, J., Rogers, T., Allesina, S., 2016. Modularity and stability in ecological communities. Nature Comm. 7, 12031. doi:10.1038/ncomms12031.

Jiang, L.-L., Perc, M., 2013. Spreading of cooperative behaviour across interdependent groups. Sci. Rep. 3 (1), 2483. doi:10.1038/srep02483.

Karrer, B., Newman, M.E.J., 2011. Stochastic blockmodels and community structure in networks. Phys. Rev. E. 83 (1), 016107. doi:10.1103/physreve.83.016107.

Kermack, W.O., McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proceed. Royal Soc. A: Math.Phys. Eng. Sci. 115 (772), 700–721. doi:10.1098/rspa.1927.0118.

Lieberman, E., Hauert, C., Nowak, M., 2005. Evolutionary dynamics on graphs. Nature 433, 312–316. doi:10.1038/nature03204.

Lozano, S., Arenas, A., Sánchez, A., 2008. Mesoscopic structure conditions the emergence of cooperation on social networks. PLOS ONE 3 (4), e1892. doi:10.1371/journal.pone.0001892.

Marcoux, M., Lusseau, D., 2013. Network modularity promotes cooperation. J. Theor. Biol. 324, 103–108. doi:10.1016/j.jtbi.2012.12.012.

May, R.M., 1972. Will a large complex system be stable? Nature 238 (5364), 413–414. doi:10.1038/238413a0.

Newman, M.E.J., 2002. Spread of epidemic disease on networks. Phys. Rev. E. 66 (1). doi:10.1103/physreve.66.016128.

Newman, M.E.J., 2006. Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E. 74 (3). doi:10.1103/physreve.74.036104.

Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. Phys. Rev. E. 69 (2), 026113. doi:10.1103/physreve.69.026113.

Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314 (5805), 1560–1563. doi:10.1126/science.1133755.

Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature 359 (6398), 826–829. doi:10.1038/359826a0.

Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M., 2006. A simple rule for the evolution of cooperation on graphs and social networks. Nature 411, 502–505. doi:10.1038/nature04605.

Ravasz, E., 2002. Hierarchical organization of modularity in metabolic networks. Science 297 (5586), 1551–1555. doi:10.1126/science.1073374.

Reichardt, J., Bornholdt, S., 2006. Statistical mechanics of community detection. Phys. Rev. E. 74 (1). doi:10.1103/physreve.74.016110.

Sah, P., Leu, S.T., Cross, P.C., Hudson, P.J., Bansal, S., 2017. Unraveling the disease consequences and mechanisms of modular structure in animal social networks. Proceed. Nation. Acad. Sci. 114 (16), 4165–4170. doi:10.1073/pnas.1613616114.

Santos, F., Rodrigues, J., Pacheco, J., 2006. Graph topology plays a determinant role in the evolution of cooperation. Proceed. Royal Soc. B. Biol.Sci. 273 (1582), 51–55. doi:10.1098/rspb.2005.3272.

Stouffer, D., Bascompte, J., 2011. Compartmentalization increases food-web persistence. PNAS 108 (9), 3648–3652. doi:10.1073/pnas.1014353108.

Trajanovski, S., Kuipers, F., Martín-Hernández, J., Mieghem, P.V., 2013. Generating graphs that approach a prescribed modularity. Comp. Comm. 36 (4), 363–372. doi:10.1016/j.comcom.2012.10.004.

Weisstein, E. W.,. Kermack-mckendrick model - from mathworld—a wolfram web resource.